

VOICE SYNTHESIZER

Publication number: JP63199399 (A)

Publication date: 1988-08-17

Inventor(s): SAKURAI ATSUSHI; TAMURA JUNICHI

Applicant(s): CANON KK

Classification:

- international: **G10L21/04; G10L21/00; (IPC1-7): G10L3/02**

- European:

Application number: JP19870031581 19870216

Priority number(s): JP19870031581 19870216

Abstract not available for **JP 63199399 (A)**

Data supplied from the *esp@cenet* database — Worldwide

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑫ 公開特許公報(A)

昭63-199399

⑪ Int. Cl.⁴
G 10 L 3/02

識別記号 庁内整理番号
A-8622-5D

⑬ 公開 昭和63年(1988)8月17日

審査請求 未請求 発明の数 2 (全18頁)

⑭ 発明の名称 音声合成装置

⑯ 特 願 昭62-31581

⑰ 出 願 昭62(1987)2月16日

⑱ 発 明 者 桜 井 稔 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
⑲ 発 明 者 田 村 純 一 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
⑳ 出 願 人 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
㉑ 代 理 人 弁理士 大塚 康徳 外1名

明 細 書

1. 発明の名称

音声合成装置

2. 特許請求の範囲

(1) 音声合成する特徴パラメータを関引きし又は重複使用することにより発声速度を変える音声合成装置において、所定時間長の音声に対応する特徴パラメータと少なくとも前記特徴パラメータ毎に対応させた発声速度制御の可否情報を記憶する記憶手段と、音声合成の際に、前記可否情報の内容が速度制御可である特徴パラメータのみを対象としてその特徴パラメータを関引きし又は重複使用する速度制御手段を備えることを特徴とする音声合成装置。

(2) 音声合成する特徴パラメータを関引きし又は重複使用することにより発声速度を変える音声

合成装置において、所定時間長の音声に対応する特徴パラメータと少なくとも前記特徴パラメータ毎に対応させた発声速度制御可否の多値情報を記憶する記憶手段と、発声速度に応じて閾値を設定する閾値設定手段と、音声合成の際に、前記多値情報の内容が前記閾値より小さい特徴パラメータのみを対象としてその特徴パラメータを関引きし又は重複使用する速度制御手段を備えることを特徴とする音声合成装置。

(3) 記憶手段は破裂性子音の破裂時点を示す特徴パラメータに対応して最大の多値情報を記憶し、続く特徴パラメータに対応して減少するような多値情報を記憶することを特徴とする特許請求の範囲第2項記載の音声合成装置。

(4) 速度制御手段は多値情報が所定の符号を有するときは無条件でその特徴パラメータの重複使

特開昭63-199399 (2)

用を行なわないことを特徴とする特許請求の範囲第2項記載の音声合成装置。

(5) 閾値設定手段は発声速度が標準速度より速いか又は遅くなるほど高い閾値を設定することを特徴とする特許請求の範囲第2項記載の音声合成装置。

3. 発明の詳細な説明

[産業上の利用分野]

本発明は音声合成装置に関し、特に音声合成する特徴パラメータを間引きし又は重複使用することにより発声速度を変える音声合成装置に関する。

[従来の技術]

音声信号は一定時間内であるとほぼ定常的である。従来は、この点に着目し、音声信号を一定時間長毎に分析し、分析結果に基づいて各区間を一組の特徴パラメータで表現せしめ、予めこれらを記憶し、音声合成の際は、これらの特徴パラメータを一定時間長毎に取り出し、順次に合成する方法が知られている。この方法は、合成操作が極めて簡単であり、音質劣化が少ないので実用に値する。具体的には、一組の特徴パラメータは一定時

3

間長の音声に対応する。従つて、特徴パラメータの組を適当に間引きし又は重複使用することにより合成音声の持続時間を増減できる。そして、従来は、この方法で発声速度を変えることが試みられていた。しかし、破裂性子音(k, t, p, b, d, g, r等)の持続時間は短いのので、たかだか1組か2組の特徴パラメータが合成されるのみである。従つて、従来方法では、間引いたり重複使用する特徴パラメータの組がたまたま破裂性子音に該当する場合には音声の明瞭度を著しく損なっていた。

[発明が解決しようとする問題点]

本発明は上述の従来技術の欠点を除去するものであり、その目的とする所は、発声速度を変えても合成音声の明瞭度を損なわない音声合成装置を提供することにある。

4

[問題点を解決するための手段]

本発明の音声合成装置は上記の目的を達成するために、所定時間長の音声に対応する特徴パラメータと少なくとも前記特徴パラメータ毎に対応させた発声速度制御の可否情報を記憶する記憶手段と、音声合成の際に、前記可否情報の内容が速度制御可能である特徴パラメータのみを対象としてその特徴パラメータを間引きし又は重複使用する速度制御手段を備えることをその概要とする。

また本発明の音声合成装置は上記の目的を達成するために、所定時間長の音声に対応する特徴パラメータと少なくとも前記特徴パラメータ毎に対応させた発声速度制御可否の多値情報を記憶する記憶手段と、発声速度に応じて閾値を設定する閾値設定手段と、音声合成の際に、前記多値情報の内容が前記閾値より小さい特徴パラメータのみを

対象としてその特徴パラメータを間引きし又は重複使用する速度制御手段を備えることをその一態様とする。

また好ましくは、記憶手段は破裂性子音の破裂時点を示す特徴パラメータに対応して最大の多値情報を記憶し、続く特徴パラメータに対応して減少するような多値情報を記憶することをその一態様とする。

また好ましくは、速度制御手段は多値情報が所定の符号を有するときは無条件でその特徴パラメータの重複使用を行わないことをその一態様とする。

また好ましくは、閾値設定手段は発声速度が標準速度より速いか又は遅くなるほど高い閾値を設定することをその一態様とする。

【作用】

7

からの発声速度指令)に応じて閾値を設定する。好ましくは、閾値設定手段は発声速度が標準速度より速いか又は遅くなるほど高い閾値を設定する。速度制御手段は、音声合成の際に、前記多値情報の内容が前記閾値より小さい特徴パラメータのみを対象としてその特徴パラメータを間引きし又は重複使用する。好ましくは、速度制御手段は多値情報が所定の符号を有するときは無条件でその特徴パラメータの重複使用を行わない。

【実施例の説明】

以下添付図面に従って本発明の実施例を詳細に説明する。

【第1実施例】

第1図は本発明による第1実施例の音声合成装置のブロック構成図である。図において、1は入力端子であり、図示せぬホスト側から送られる発

特開昭63-199399(3)

かかる構成において、記憶手段は所定時間長の音声に対応する特徴パラメータと少なくとも前記特徴パラメータ毎に対応させた発声速度制御の可否情報(例えば2値情報)を記憶する。速度制御手段は、音声合成の際に、前記可否情報の内容が速度制御可である特徴パラメータのみを対象としてその特徴パラメータを間引きし又は重複使用する。

またかかる構成において、記憶手段は所定時間長の音声に対応する特徴パラメータと少なくとも前記特徴パラメータ毎に対応させた発声速度制御可否の多値情報を記憶する。好ましくは、記憶手段は破裂性子音の破裂時点を示す特徴パラメータに対応して最大の多値情報を記憶し、続く特徴パラメータに対応して減少するような多値情報を記憶する。閾値設定手段は発声速度(例えば外部

8

声指令及び発声速度指令等を入力する。2は中央演算装置(CPU)であり、入力した発声指令及び発声速度指令に従って合成音声の発声及び速度制御を行なう。2AはCPU2が実行する制御プログラムを記憶しているメモリ(ROM)であり、例えば第6図に示す第1実施例の制御プログラム又は第10図に示す第2実施例の制御プログラムを記憶している。更に、3は速度制御の可否情報と共に音声の特徴パラメータの組を収納している第1記憶装置、4はCPU2が使用する補助記憶装置、5はPARCOR型音声合成器、6はD/A変換器、7は増幅器、8は音声出力用のスピーカである。

第2図(A)～(C)は同一男性の発声した「ミタイ」の一部「タイ」の音声波形を示す図に係り、第2図(A)は丁寧に発声した場合の音声

特開昭63-199399 (4)

波形、第2図(B)は約1.5倍の速さで発声した場合の音声波形、第2図(c)は約2倍の速さで発声した場合の音声波形を示している。

第3図(A)～(C)は第2図(A)～(C)の各音声波形の一部を時間軸方向に同一倍率で拡大した図に係り、音声「タ」の開始部分を示している。音声波形の下目盛は1目盛が長さ10ミリ秒のフレームであり、各フレームはその区間の音声波形を一組の特徴パラメータで表現する。例えば、第3図(A)のフレーム(a₁)は子音「t」の破裂時点の特徴を示している。これを、第3図(B)のフレーム(b₁)又は第3図(C)のフレーム(c₁)と比較すれば解るように、子音「t」の破裂時点の特徴は発声速度が変化しても殆ど変っていない。従つて、逆に発声速度を変化させる場合は、もし破裂時点の特徴フ

レームに対して間引きや重複使用が行われると、特徴が著しく変化し、音声の明瞭度を損なう。この点は、他の破裂性子音(k, p, d, g, r等)の場合も同じである。そこで、第1実施例では音声波形をフレーム単位で分析した特徴パラメータの形で格納する際に、速度制御の可否情報をフレーム毎に付加し、例えば破裂性子音の破裂時点のフレームのように、間引きや重複使用の対象とすべきでないフレームに対しては可否情報の内容を“否”とする。

第4図は第1実施例の可否情報及び特徴パラメータの組の構造を示す図である。男声「ミタイ」を各10ミリ秒のフレームで分析すると、特徴パラメータの組の総フレーム数は1～NのN個である。そして各フレームにおける特徴パラメータの組は夫々ビットP_i(iはフレーム番号)、アン

1 1

プA_i及びPARCOR係数K_iから成る。またフレーム毎に速度制御の可否情報eを付している。可否情報eの内容が“0”のときは速度制御(間引き、重複使用)可であり、“1”のときは速度制御不可である。

第5図(A)は第1実施例における速度指令vとフレームの間引き又は重複使用の周期mとの関係を示す図である。図において、速度指令vの内容は標準速度のときに“0”とする。この場合、CPU2は第4図の特徴パラメータの組を全部そのまま出力する。標準より速い速度指令vは正の整数“1～4”で表わす。この場合、CPU2は演算 $m = 6 - |v|$ を実行して周期mを求め、かつ速度指令vの符号は正であるから周期m毎に間引き可否の制御を行う。即ち、mフレーム毎に速度制御の可否情報eの内容を調べ、もしeの内容

1 2

が“0”(可)であればそのフレームの特徴パラメータの組のPARCOR型音声合成器5への転送の間引き。標準より遅い速度指令vは負の整数“-1～-4”で表わす。この場合、CPU2は演算 $m = 6 - |v|$ を実行して周期mを求め、かつ速度指令vの符号は負であるから周期m毎に重複使用可否の制御を行う。即ち、mフレーム毎に速度制御の可否情報eの内容を調べ、もしeの内容が“0”(可)であればそのフレームの特徴パラメータの組を重複使用してPARCOR型音声合成器5に転送する。

第6図は第1実施例の速度制御手順を示すフローチャートである。第1図の入力端子1から発声指令及び速度指令vが入力されると第6図の処理を開始する。第6図において、変数jはフレームの計数値(フレーム番号)を示しており、1～

N の値をとる。変数（周期カウンタ） n は間引き又は重複使用をするための周期 m を計数しており、 $0 \sim m-1$ の値をとる。フラグ f は 1 周期内での間引き又は重複使用の処理の完了状態を示しており、1 周期の開始時点では周期カウンタ n と共に“0”にリセットされ、間引き又は重複使用を行うと“1”にセットされる。またフラグ f は、重複処理の際は、同一の特徴パラメータを 2 回使用する指標として 1 次的に数値“-1”がセットされる。

<初期処理>

ステップ S1 では演算 $m = 6 - |v|$ を行つて周期 m を求める。ステップ S2 ではフレーム番号 j に数値 1 をセットしてフレーム (1) からの音声パラメータ（可否情報 e を含む）のアクセスを可能にする。ステップ S3 では周期カウンタ n と

フラグ f をリセットする。ステップ S4 では速度指令 v の内容を読む。

<標準速度の発声>

ステップ S4 の判別で速度指令 v の内容が“0”のときは標準速度の発声である。フローはステップ S11 に進み、当該フレーム j の特徴パラメータの組を P A R C O R 型音声合成器 5 に転送する。更に、P A R C O R 型音声合成装置 5 は転送された特徴パラメータの組を音声情報に合成し、D/A 変換器 6 は合成された音声情報をアナログ信号に変換し、増幅器 7 はアナログ信号を増幅し、スピーカ 8 は合成音声を出力する。

一方、CPU 2 はこの間に、ステップ S12 ではフラグ f を調べ、“-1”ではないからステップ S13 に進み、フレーム番号 j に +1 する。ステップ S15 では略 1 フレームの時間長（10 ミ

特開昭63-199399 (5)

1 5

リ秒)だけ時間を持ち、ステップ S16 ではフレーム番号 j が総フレーム数 N に達したか否かを調べる。もし N に達していれば総フレーム数 N の出力を完了したので、ステップ S19 に進み、処理を終了する。また N に達していない間は、ステップ S17 に進み、周期カウンタ n に +1 する。ステップ S18 では $n < m$ か否かを調べ、もし $n < m$ なら 1 周期の途中であるからステップ S4 に戻り、次のフレームの読み出しを行う。また、 $n < m$ でないなら $n = m$ (次の周期の始まり)であるから、ステップ S3 に戻り、周期カウンタ n 及びフラグ f をリセットする。こうして、標準速度の発声では無条件で総フレーム数 N の特徴パラメータの組が出力される。

<標準速度よりも速い発声>

ステップ S4 の判別で速度指令 v の内容が

1 6

“0”でないときは標準速度よりも速いか、遅い発声である。そして速度指令 v の符号が正のときは標準速度よりも速い発声であり、以下の間引き処理を行う。さて、ステップ S3 を通り、周期カウンタ $n = 0$ 及びフラグ $f = 0$ のタイミング（1 周期の始め）は間引き処理の可否を調べるタイミングである。ステップ S5 ではフラグ f の内容を読む。フラグ f の内容は始めは“0”であるからステップ S6 に進み、当該フレームの速度制御の可否情報 e_1 を読み出す。ステップ S7 では e_1 の内容が“0”か否かを調べる。もし可否情報 e_1 の内容が“0”なら当該フレームは速度制御を可とされたフレームであり、フローはステップ S8 に進み、速度指令 v の符号が正か否かを調べる。今、速度指令 v の符号は正であるから、ステップ S10 に進み、フラグ f に“1”をセット

1 7

---789---

1 8

特開昭63-199399(6)

して間引き処理完了した旨の置置をする。ステップS12の判別では、フラグfの内容は真でないからステップS13に進み、フレーム番号jに+1する。こうして間引きの処理は、ステップS11の処理を行わずにフレーム番号を1つ更新することにより完了する。ステップS15では所定の時間待ちをする。この場合の待ち時間は1フレーム時間長ではない。そして、次にステップS5に戻ったときはフラグfの内容は“0”でない。即ち、これ以降の1周期内では常にステップS5からステップS11に進み、標準速度の発声において述べたと同様にして順々に特徴パラメータの組を読み出し、P A R C O R型音声合成器5に転送する。かようにして、周期m毎にその最初のフレームの可否情報e_jの内容が調べられ、もし“可”ならばそのフレームの特徴パラ

メータの組が間引かれる。

しかし、ステップS7の判別において可否情報e_jの内容が“1”のときは当該フレームの特徴パラメータの組の転送を間引かない。フローはステップS11に進み、当該フレームの特徴パラメータの組をP A R C O R型音声合成器5に転送する。従つて、このフレーム処理ではフラグfに“1”がセットされないから、次のステップS5の判別でもフラグf=0を満足する。そして、ステップS5では次のフレームの可否情報e_{j+1}の内容が調べられ、もしこれが“0”のときはこのフレームについて間引きの処理が行なわれる。このようにして、標準速度よりも速い発声の場合は、周期m毎に間引きの制御が実行され、もし当該フレームの特徴パラメータの組を間引けないときはその次のフレームが可否情報eの内容に従つて間

19

引かれることにより、標準速度よりも速い発声を常に忠実に達成し、しかも重要な破裂性子音の破裂時点のフレームは失われない。

<標準速度よりも速い発声>

ステップS8の判別で速度指令vの符号が負のときは標準速度よりも速い発声の場合であり、以下の重複使用の処理を行う。同様にして、ステップS3を通り、周期カウンタn=0及びフラグf=0のタイミングは重複使用の処理の可否を調べるタイミングである。ステップS5ではフラグfの内容を調べる。フラグfの内容は始めは“0”であるからステップS6に進み、当該フレームの速度制御の可否情報e_jを読み出す。ステップS7ではe_jの内容が“0”か否かを調べる。もし可否情報e_jの内容が“0”なら当該フレームは速度制御を可とされたフレームであ

20

り、フローはステップS8に進み、速度指令vの符号が正か否かを調べる。今、速度指令vの符号は負であるから、ステップS9に進み、同一の特徴パラメータを2回使用する指標として1次的にフラグfに数値“-1”をセットする。ステップS11では1回目の特徴パラメータの組をP A R C O R型音声合成器5に転送する。ステップS12の判別ではフラグfの内容が“-1”であることによりステップS14に進み、フラグfに“1”をセットする。重複使用のため特徴パラメータの組を1回余分に転送完了した旨の置置である。またステップS13の処理をスキップすることによりフレーム番号を更新しない。即ち、このフレーム番号の特徴パラメータを2度使用する。こうして、次からのステップS5の判別においてはフラグfの内容が“1”であることにより、当

21

22

特開昭63-199399(7)

該1周期を完了するまではフレーム番号Jを更新して各特徴パラメータの組をP A R C O R型音声合成器5に転送する。

しかし、ステップS7の判別において可否情報eの内容が“1”のときは当該フレームの特徴パラメータの組の重複転送を行わない。フローはステップS11に進み、当該フレームの特徴パラメータの組をP A R C O R型音声合成器5に転送し、ステップS13でフレーム番号を1つ更新する。従つて、このフレーム処理ではフラグfに“1”がセットされないから、次のステップS5の判別でもフラグf=0を満足する。そして、ステップS6では次のフレームの可否情報e₁の内容が調べられ、もしこれが“0”のときはこのフレームについて重複使用の処理が行なわれる。このようにして、標準速度よりも遅い発声の場合

は、周期m倍に重複使用の制御が実行され、もし当該フレームの特徴パラメータの組を重複使用できないときはその次のフレームの可否情報eの内容に従つて次のフレームの特徴パラメータの組を重複使用することにより、標準速度よりも遅い発声を常に忠実に達成し、しかも重要な破裂性子音の破裂時点のフレームは重複使用されない。

第7図(A)は第1実施例の可否情報eに対する4種類の速度指令vにおける処理結果を示した図に係り、音声「タ」の開始部分のフレーム(i-2)からフレーム(i+5)までの8フレームについての各処理結果を示している。図において、“×”印は間引きしたフレームを表し、“◎”印は重複使用したフレームを表す。またフレーム(i-2)の位置は何れの速度でも丁度1周期mの倍数の位置にあると仮定する。

2 3

速度指令v=2の場合は、演算 $m=8-12$ により、周期 $m=4$ である。従つて、先頭フレーム(i-2)と、それから4つ目のフレーム(i+2)の可否情報eの内容が調べられ、この場合は何れも“0”(可)であるので、共に間引かれる。

速度指令v=3の場合は、演算 $m=8-13$ により周期 $m=3$ である。従つて、先頭フレーム(i-2)と、それから3つ目のフレーム(i+1)等の可否情報eの内容が調べられ、この場合は何れも“0”であるので、共に間引かれる。

速度指令v=4の場合は、演算 $m=8-14$ により周期 $m=2$ である。従つて、先頭フレーム(i-2)と、2つ目のフレーム(i)と、更に2つ目のフレーム(i+2)等の可否情報eの内

2 4

容が調べられる。この場合はフレーム(i-2)及びフレーム(i+2)については何れも“0”であるので、共に間引かれる。しかし、フレーム(i)については可否情報eの内容が“1”(不可)であるので、当該フレームの特徴パラメータの組は間引かれずに、その次のフレーム(i+1)の可否情報eの内容が調べられ、この場合は内容が“0”であるので間引かれる。こうして、平均の発声速度には影響を与えず、しかも破裂性子音「t」の破裂時点を示すフレーム(i)の特徴パラメータは間引かれることなくそのまま合成器5に転送されるので、明瞭性のある発声が合成される。

速度指令v=-4の場合は演算 $m=8-14$ により周期 $m=2$ である。従つて、先頭フレーム(i-2)と、2つ目のフレーム(i)と、更に

2 5

---791---

2 6

2つ目のフレーム(1+2)等の可否情報eの内容が調べられる。またvの符号が負であるので、フレームの重複使用が行われる。即ち、この場合もフレーム(i-2)及びフレーム(1+2)については何れも“0”であるので、併に重複使用が行われる。しかし、フレーム(1)については可否情報eの内容が“1”(不可)であるので、当該フレームの特徴パラメータの組は重複使用されず、その次のフレーム(i+1)の可否情報eの内容が調べられ、この場合は内容が“0”であるので、重複使用される。この場合も、平均の発声速度には影響を与えず、しかも破裂性子音「t」の破裂時点を示すフレーム(i)の特徴パラメータは重複使用されることがなく、図だけ合成部に転送されるので、破裂音がダブらず、明瞭な音声合成される。

2.7

※や重複使用を速度指令vの大小に応じて適応的に行わせしめ、発声速度を変化させた場合にもより自然で、明瞭な音声合成出力することにある。

第3図(A)に戻り、今度は無声破裂性子音「t」の破裂時点のフレーム(a₁)とその次のフレーム(a₁₊₁)に着目する。前述の如く、速度指令vを変えても、破裂時点のフレーム(a₁)、(b₁)及び(c₁)についてはさほど変化が認められなかつた。しかし、次のフレーム(a₁₊₁)に着目すると、1.5倍の速さのフレーム(b₁₊₁)との間では殆ど不変であるのに対し、2倍の速さのフレーム(c₁₊₁)と比較すると、もはやフレーム(a₁₊₁)の特徴を喪失するフレームは見当らない。これは発声が速くなるのに従い、子音「t」から後続の母音「a」への長

特開昭63-199399(8)

第8図(A)は第1実施例の可否情報eに対する4種類の速度指令vにおける処理結果を原音声の波形と共に示した図に係り、音声「タ」の開始部分のフレーム(i-2)からフレーム(1+5)までの8フレームについての各処理結果が示されている。第7図(A)と同様に、“x”印は間引きしたフレームを表し、“◎”印は重複使用したフレームを表す。第8図(A)より明らかな通り、無声破裂性子音「t」の破裂時点を示すフレーム(1)の信号は、発声速度vの如何に依らず間引きや重複使用の対象とはなっていない。

〔第2実施例〕

第2実施例のブロック構成図は第1図のものと同一である。第2実施例の特徴は、第1実施例で各フレーム毎に付加した1ビットの可否情報eを多値化して利用することにより、フレームの間引き

2.8

音結合部が短くなる為であり、この点は他の破裂性子音(k, p, b, d, g, r等)の場合も同様である。

そこで、第2実施例では発声速度vを変える場合に、標準より速い発声においては、破裂時点のフレームは間引かず、かつ後続母音部への長音結合部フレームについてはその間引き法を工夫し、即ち、速度指令vの大小に応じて間引き法を適応的に変化させることにより、自然に近い音声合成出力する。また標準より遅い発声においては、破裂性子音の継続時間をある程度以上長くすると子音部の音韻性が失われることが知られているので、第2実施例では多値化した可否情報eに対してフレームの間引きだけを禁止する符号情報も付加して処理を行い、フレームの重複使用による破裂性子音の音韻性の变化を防止する。

特開昭63-199399 (9)

第9図は第2実施例の可否情報及び特徴パラメータの組の構造を示す図である。図において、フレーム番号及び特徴パラメータの組に関しては第4図のものと同一であるが、速度制御の可否情報 e_2 は異なる。第2実施例の可否情報 e_2 は図のように符号化されており、“0”を含む正又は負の整数で表わされる。

そして、可否情報 e_2 の内容は、その絶対値が速度指令 v に応じて決定された所定閾値 θ より以下のときは、当該フレームの間引きや重複使用を可とし、また所定閾値 θ より大きいときは当該フレームの特徴パラメータの組をそのまま音声出力させるように利用される。

また可否情報 e_2 の内容に負の符号が付されたときは、常に重複使用の対象から外される。即ち、速度指令 v が標準より速い場合はその可否

情報 e_2 の内容が負でないフレームのみを対象として上記の処理を行う。

第9図において、無声破裂性子音「 t 」の破裂時点のフレーム(1)に対しては例えば最大の絶対値「8」を与え、以下の後続の母音「 e 」に至る異音結合部の3フレームに対しては夫々絶対値「3」、「2」、「1」を与えている。このような傾斜特性を与えると、速度指令 v が標準速度に近い時(閾値 θ が低い時)は母音定常部に近いフレームのみが間引きや重複使用の対象となり、速度指令 v が標準速度から外れる(閾値 θ が高くなる)に従って破裂時点に近いフレームまで間引きや重複使用の対象になる。またその際に、子音部分のフレーム(1)及びフレーム(1+1)には負の符号を与え、重複使用を無条件に禁止して音質変化を防止している。

3.1

第5図(8)は第2実施例における発声速度 v 、閾値 θ 及び間引き又は重複使用の周期 m の関係を示す図である。同様に、速度指令 v の内容は標準の発声速度を“0”とし、標準速度より速い場合を正の整数“1~4”で表わし、標準速度より遅い場合を負の整数“-1~-4”で表わしている。そして、閾値 θ 及び周期 m の値は速度指令 v の内容を用いて下記の演算(1)及び(2)により決定する。

$$\theta = |v| - 1 \quad \cdots (1)$$

$$m = 8 - |v| \quad \cdots (2)$$

従つて、もし速度指令 v が標準の“0”のときは演算(1)により閾値 $\theta = -1$ になるから、この場合は可否情報 e_2 の絶対値は閾値 θ 以下の値を取り得ない。従つて、常にフレームの間引きも重複使用も起こらず、全フレームの特徴パラメー

3.2

タの組がそのまま合成出力される。

こうして、入力端子1から発声指令及び速度指令 v が入力されると、CPU2は演算(1)及び(2)を実行して閾値 θ と周期 m を求め、もし速度指令 v の内容が“0”か正の整数であれば m フレーム毎に可否情報 e_2 の内容を調べ、その絶対値が閾値 θ 以下であるときは当該フレームの特徴パラメータの組の間引き。また速度指令 v が負の整数であれば、 m フレーム毎に可否情報 e_2 の内容を調べ、 e_2 の符号が負でなく、かつ閾値 θ 以下であるときは当該フレームの特徴パラメータの組を重複使用する。

第10図は第2実施例の速度制御手順を示すフローチャートである。尚、第8図と同一の処理には同一のステップ番号を付して説明を省略する。

特開昭63-199399(10)

<初期処理>

入力端子1から発声指令及び速度指令 v が入力されるとステップS100に入力する。ステップS100では前記の演算(1)及び(2)に従って閾値 $t = |v| - 1$ と周期 $m = 6 - |v|$ を求める。

<標準速度の発声>

ステップS101の判別で速度指令 v の内容が“0”のときは標準速度の発声である。フローはステップS105に進み、 $|e_{2j}| > t$ もか否かの判別をする。ところで、標準速度のときは閾値 $t = |0| - 1 = -1$ であるから、 $|e_{2j}| > t$ を必ず満足する。従つて、全フレームを通じてステップS105を実行し、標準速度の発声が実行される。

<標準速度よりも速い発声>

3 5

<標準速度よりも速い発声>

ステップS101の判別で速度指令 v の符号が負のときは標準速度よりも速い発声の場合であり、以下の重複使用の処理を行う。同様にして、ステップS3を通り、周期カウンタ $n = 0$ 及びフラグ $f = 0$ のタイミングは重複使用の処理の可否を調べるタイミングである。ステップS6では可否情報 e_{2j} を取り出し、ステップS102では $e_{2j} < 0$ もか否かの判別をする。 $e_{2j} < 0$ のときは重複使用禁止フレームと判断して無条件で重複使用を行わない。フローはステップS106に進み、当該フレームの特徴パラメータの組を転送し、ステップS13に進み、フレーム番号を更新する。

また $e_{2j} < 0$ でないときは閾値による制御に従う。即ち、ステップS103では $e_{2j} > t$ もか否か

ステップS101の判別で速度指令 v の内容が正の整数ときは標準速度よりも速い発声である。同様にして、ステップS3を通り、周期カウンタ $n = 0$ 及びフラグ $f = 0$ のタイミング(1周期の始め)は間引き処理の可否を調べるタイミングである。ステップS6では可否情報 e_{2j} を取り出し、ステップS105では $|e_{2j}| > t$ もか否かの判別をする。もし $|e_{2j}| > t$ を満足するときは、当該フレームを間引かないでステップS106に進む。この場合はフラグ f に“1”を立てないので、次のフレームについてもステップS105では $|e_{2j}| > t$ もか否かの判別をする。また $|e_{2j}| > t$ を満足するときは、ステップS107に進み、当該フレームを間引いて、フラグ f に“1”を立て、間引き処理完了の旨を宣言する。

3 6

の判別をする。 $e_{2j} > t$ のときは重複使用不可フレームと判断して重複使用を行わない。また $e_{2j} > t$ でないときはステップS104に進み、フラグ f に“-1”をセットして当該フレームの重複使用を可能にする。

第7図(8)は第2実施例の可否情報 e_{2j} に対する4種類の速度指令 v における処理結果を示した図に係り、音声「タ」の開始部分のフレーム($i - 2$)からフレーム($i + 5$)までの8フレームについての各処理結果を示している。同様にして、“×”印は間引きしたフレームを表わし、“⊖”印は重複使用したフレームを表わす。またフレーム($i - 2$)の位置は何れの速度においても丁度1周期 m の倍数の位置にあると仮定した。

まず速度指令 $v = 2$ の場合は、 v の内容が正で

特開昭63-199399 (11)

あるから間引き制御の対象になる。演算(1)及び演算(2)により閾値 $t=1$ 、周期 $m=4$ が求まる。従つて、各先頭フレームはフレーム $(i-2)$ 及びフレーム $(i+2)$ であり、対応する可否情報 e_2 の絶対値が閾値 t と比較される。フレーム $(i-2)$ では $|e_{2(i-2)}|=0$ で閾値 $t=1$ 以下であるから間引きの対象になる。しかし、フレーム $(i+2)$ では $|e_{2(i+2)}|=2$ で閾値 $t=1$ より大きいから間引きは行われない。次のフレーム $(i+3)$ においては、 $|e_{2(i+3)}|=1$ で閾値 $t=1$ 以下となり、間引きが行われる。こうして、速度指令 $v=2$ の場合はフレーム $(i-2)$ 及びフレーム $(i+3)$ の特選パラメータの組が間引きされる。

次に速度指令 $v=3$ の場合は、 v の内容が正であるから間引き制御の対象になる。演算(1)

及び演算(2)により閾値 $t=2$ 、周期 $m=3$ が求まる。従つて、各先頭フレームはフレーム $(i-2)$ 、フレーム $(i+1)$ 及びフレーム $(i+4)$ である。フレーム $(i-2)$ 及びフレーム $(i+4)$ では夫々 $|e_{2(i-2)}|=|e_{2(i+4)}|=0$ であり、閾値 $t=2$ 以下であるから間引きの対象になる。しかし、フレーム $(i+1)$ では $|e_{2(i+1)}|=3$ で閾値 $t=2$ より大きいから、間引きは行われない。次のフレーム $(i+2)$ では、 $|e_{2(i+2)}|=2$ で閾値 $t=2$ 以下となり、間引きが行われる。こうして、速度指令 $v=3$ の場合はフレーム $(i-2)$ 、フレーム $(i+2)$ 及びフレーム $(i+4)$ の特選パラメータの組が間引きされる。

次に速度指令 $v=4$ の場合は、 v の内容が正であるから間引き制御の対象になる。演算(1)

3 9

及び演算(2)により、閾値 $t=3$ 、周期 $m=2$ が求まる。従つて、各先頭フレームはフレーム $(i-2)$ 、フレーム (i) 、フレーム $(i+2)$ 及びフレーム $(i+4)$ である。フレーム $(i-2)$ 、フレーム $(i+2)$ 及びフレーム $(i+4)$ においては $|e_{2(i-2)}|=0$ 、 $|e_{2(i+2)}|=2$ 、 $|e_{2(i+4)}|=0$ であり、何れも閾値 $t=3$ 以下であるから間引きの対象になる。しかし、フレーム (i) では $|e_{2(i)}|=8$ で閾値 $t=3$ より大きいから間引きは行われない。次のフレーム $(i+1)$ においては、 $|e_{2(i+1)}|=3$ で閾値 $t=3$ 以下となり、間引きが行われる。こうして、速度指令 $v=4$ においてはフレーム $(i-2)$ 、フレーム $(i+1)$ 、フレーム $(i+3)$ 及びフレーム $(i+4)$ の特選パラメータの組が間引きされる。

4 9

最後に速度指令 $v=-4$ の場合は、 v の内容が負であるから重複使用の制御対象である。演算(1)及び演算(2)により閾値 $t=3$ 、周期 $m=2$ が求まる。従つて、各先頭フレームはフレーム $(i-2)$ 、フレーム (i) 、フレーム $(i+2)$ 及びフレーム $(i+4)$ である。そこで、各対応する可否情報 e_2 の値が調べられ、負でなければ閾値 t と比較される。フレーム $(i-2)$ 、フレーム $(i+2)$ 及びフレーム $(i+4)$ では可否情報 e_2 の値が負でない。そして、 $|e_{2(i-2)}|=0$ 、 $|e_{2(i+2)}|=2$ 、 $|e_{2(i+4)}|=0$ であるから何れも閾値 $t=3$ 以下であり、重複使用の対象となる。しかし、フレーム (i) では $e_{2(i)}=-8$ であるから、可否情報 e_2 の値が負であるので重複使用は行われない。また、同1周期内の残りのフレーム

($i+1$)でも $e_{x(i)}=-3$ で、負であるから重複使用は行われない。こうして、速度指令 $v=-4$ においてはフレーム($i-2$)、フレーム($i+2$)及びフレーム($i+4$)で重複使用される。

第8図(B)は第2実施例の可否情報 e_x に対する4種類の速度指令 v における処理結果を原音声の波形と共に示した図に係り、音声「タ」の開始部分のフレーム($i-2$)からフレーム($i+5$)までの8フレームについての各処理結果が示されている。第7図(B)と同様に、“×”印は間引きしたフレームを表わし、“○”印は重複使用したフレームを表わす。第8図(B)より明らかな通り、標準速度より速い発声においては、無声破裂性子音「も」の破裂時点のフレームが常に保存され、更に破裂時点と母音

特開昭63-199399(12)

「a」の定常部とを結合調音結合部分のフレーム($i+1$)、フレーム($i+2$)及びフレーム($i+3$)が発声速度 v の増加に応じて、母音定常部に近いフレームから順に適応的に間引きされているので、合成出力される音声は発声速度 v に依らず、その明瞭性及び自然性を保つことができる。また、標準速度より遅い発声においては、無声破裂性子音「も」の特徴を示すフレーム(i)及びフレーム($i+1$)が重複使用されずにそのまま転送されているので、合成出力される音声は破裂性の子音部が時間軸方向に延長されず、その音韻性を保つことができる。

尚、上述の実施例では音声の特徴を表す特徴パラメータ及び音声合成器としてP A R C O R係数及びP A R C O R型音声合成器を用いたが、1定時間長の音声をも1組のパラメータで表現する合成

4 3

方式であれば、いかなる方式でも実施可能であることは明白である。

また、第2実施例において、間引きや重複使用の閾値 θ を速度指令 v を定数とする1次式の形で与えたが、速度指令 v 毎に独立した手続で与えることが出来ることは明白である。

更に、第2実施例において、破裂性子音部と後続母音部定常部への調音結合部分について速度制御の可否情報 e_x の効果を説明したが、本発明の効果はそれに限定されず、合成出力する音声の如何なる部分にも適用可能なことは明らかである。

〔発明の効果〕

以上述べた如く、従来は単に機械的に行われていた特徴パラメータの間引きや重複使用を、本発明によれば、速度制御の可否情報を負荷し、速度

4 4

指令 v の大小に応じて特徴パラメータの間引きや重複使用を適応的に行うため、音調変化や脱落のない明瞭で且つ自然性を持った音声を合成することができる。

4. 図面の簡単な説明

第1図は本発明による第1実施例の音声合成装置のブロック構成図。

第2図(A)～(C)は同一男性の発声した「ミタイ」の一部「タイ」の音声波形を示す図。

第3図(A)～(C)は第2図(A)～(C)の各音声波形の一部を時間軸方向に同一倍率で拡大した図。

第4図は第1実施例の可否情報及び特徴パラメータの組の構造を示す図。

第5図(A)は第1実施例における速度指令 v

特開昭63-199399 (13)

とフレームの間引き又は重複使用の周期 m との関係を示す図、

第5図(B)は第2実施例における発声速度 v 、間値 m 及び間引き又は重複使用の周期 m の関係を示す図、

第6図は第1実施例の速度制御手順を示すフローチャート、

第7図(A)は第1実施例の可否情報 a に対する4種類の速度指令 v における処理結果を示した図、

第7図(B)は第2実施例の可否情報 a に対する4種類の速度指令 v における処理結果を示した図、

第8図(A)は第1実施例の可否情報 a に対する4種類の速度指令 v における処理結果を原音声の波形と共に示した図、

第8図(B)は第2実施例の可否情報 a に対する4種類の速度指令 v における処理結果を原音声の波形と共に示した図、

第9図は第2実施例の可否情報及び特徴パラメータの組の構造を示す図、

第10図は第2実施例の速度制御手順を示すフローチャートである、

図中、1…入力端子、2…中央演算装置(CPU)、3…第1記憶装置、4…補助記憶装置、5…PARCOR型音声合成器、6…D/A変換器、7…増幅器、8…スピーカである。

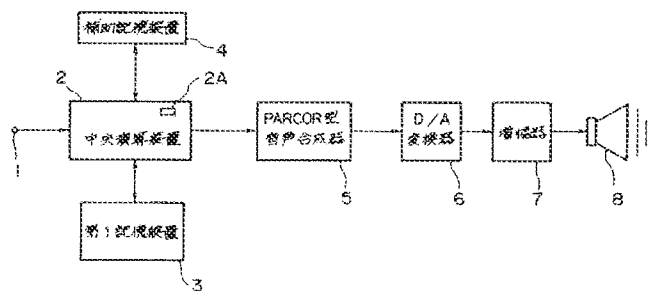
特許出願人 キヤノン株式会社

代理人 井理士 大塚 廣

(他1名)

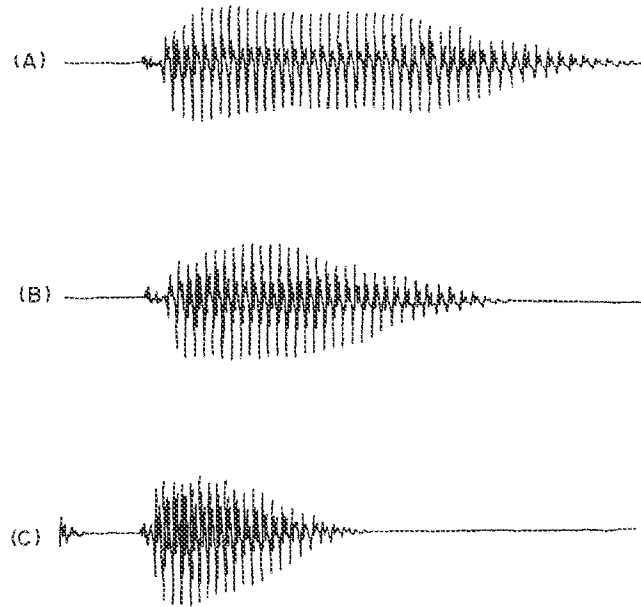
47

48

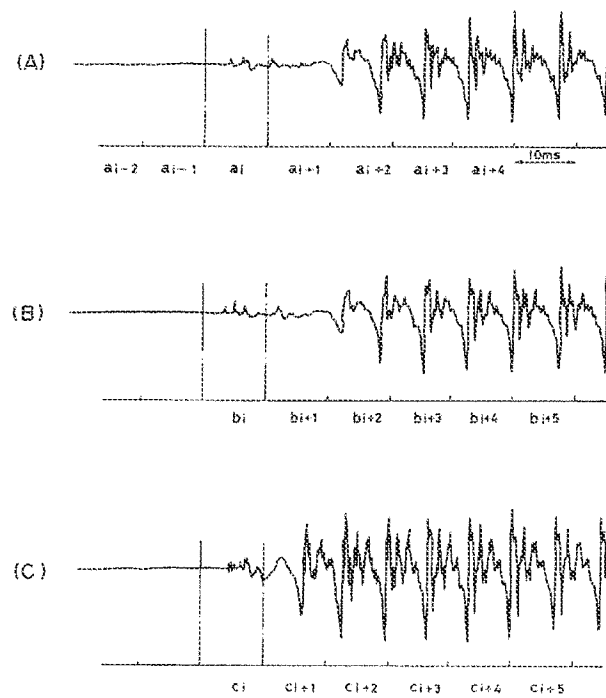


第 1 図

特開昭63-199399 (14)



第 2 図



第 3 図

特開昭63-199399 (15)

| フレーム | 可変情報 e | 呼称パラメータ (エンタナフ/PARCOR等式) |
|-------|-----------|-----------------------------------|
| 1 | 0 | $P_1, A_1, K_1 \dots$ |
| 2 | 0 | $P_2, A_2, K_2 \dots$ |
| <hr/> | | |
| $i-3$ | 0 | $P_{i-3}, A_{i-3}, K_{i-3} \dots$ |
| $i-2$ | 0 | $P_{i-2}, A_{i-2}, K_{i-2} \dots$ |
| $i-1$ | 0 | $P_{i-1}, A_{i-1}, K_{i-1} \dots$ |
| i | 1 | $P_i, A_i, K_i \dots$ |
| $i+1$ | 0 | $P_{i+1}, A_{i+1}, K_{i+1} \dots$ |
| $i+2$ | 0 | $P_{i+2}, A_{i+2}, K_{i+2} \dots$ |
| $i+3$ | 0 | $P_{i+3}, A_{i+3}, K_{i+3} \dots$ |
| $i+4$ | 0 | $P_{i+4}, A_{i+4}, K_{i+4} \dots$ |
| $i+5$ | 0 | $P_{i+5}, A_{i+5}, K_{i+5} \dots$ |
| <hr/> | | |
| $N-1$ | 0 | $P_{N-1}, A_{N-1}, K_{N-1} \dots$ |
| N | 0 | $P_N, A_N, K_N \dots$ |

第 4 図

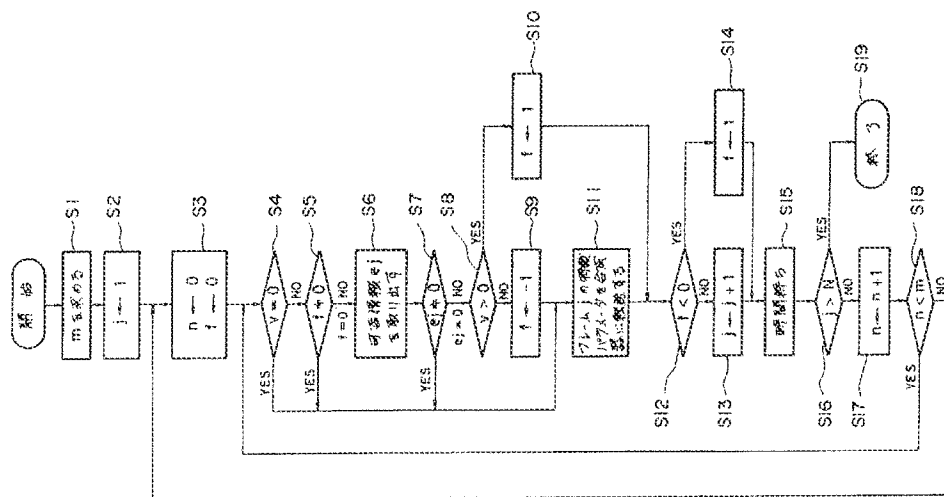
| 発声の差 (v) | 間引き又は重複の 間隔 (m) | |
|-------------|--------------------|-------|
| -4 | 2 | 重 複 |
| -3 | 3 | |
| -2 | 4 | |
| -1 | 5 | |
| 0 | 6 | 間 引 き |
| 1 | 5 | |
| 2 | 4 | |
| 3 | 3 | |
| 4 | 2 | |

(A)

| 発声の差 (v) | 間 隔 (t) | 間引き又は重複の 間隔 (m) |
|-------------|------------|--------------------|
| -4 | 3 | 2 |
| -3 | 2 | 3 |
| -2 | 1 | 4 |
| -1 | 0 | 5 |
| 0 | -1 | 6 |
| 1 | 0 | 5 |
| 2 | 1 | 4 |
| 3 | 2 | 3 |
| 4 | 3 | 2 |

(B)

第 5 図



第 6 図

(A)

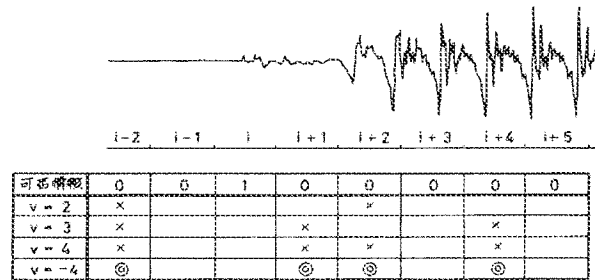
| フレーム | 可変情報 e_i | $y=2$ | $y=3$ | $y=4$ | $y=-4$ |
|-------|---------------|-------|-------|-------|--------|
| $i-2$ | 0 | x | | | ⊖ |
| $i-1$ | 0 | | | | |
| i | 1 | | | | |
| $i+1$ | 0 | | x | x | ⊖ |
| $i+2$ | 0 | x | | x | ⊖ |
| $i+3$ | 0 | | | | |
| $i+4$ | 0 | | x | x | ⊖ |
| $i+5$ | 0 | | | | |

(B)

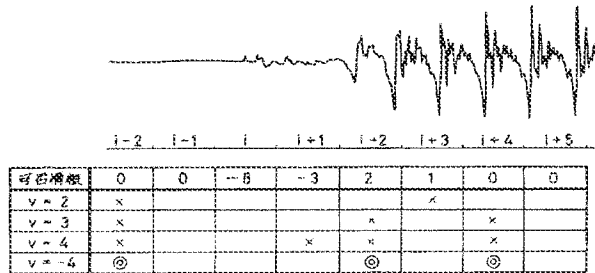
| フレーム | 可変情報 e_i | $y=2$ | $y=3$ | $y=4$ | $y=-4$ |
|-------|---------------|-------|-------|-------|--------|
| $i-2$ | 0 | x | x | x | ⊖ |
| $i-1$ | 0 | | | | |
| i | -8 | | | | |
| $i+1$ | -3 | | | x | |
| $i+2$ | 2 | | x | x | ⊖ |
| $i+3$ | 1 | x | | | |
| $i+4$ | 0 | | x | x | ⊖ |
| $i+5$ | 0 | | | | |

第 7 図

特開昭63-199399 (17)



(A)



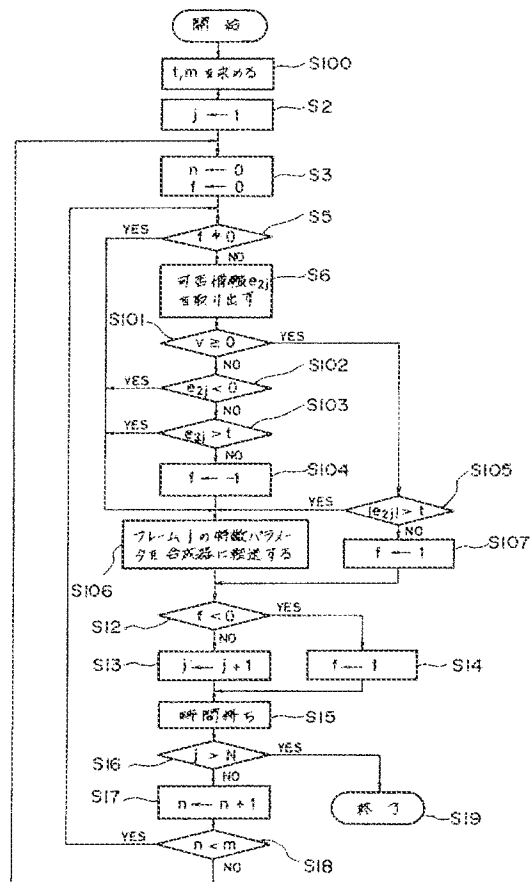
(B)

第 8 図

| フレーム | 可否情報 (e ₂) | 特徴パラメータ (ピーチング, PARCOR係数) |
|-------|---------------------------|---|
| 1 | 3 | P ₁ , A ₁ , K ₁₁ |
| 2 | 2 | P ₂ , A ₂ , K ₂₁ |
| <hr/> | | |
| $i-3$ | 0 | P _{i-3} , A _{i-3} , K _{i-31} |
| $i-2$ | 0 | P _{i-2} , A _{i-2} , K _{i-21} |
| $i-1$ | 0 | P _{i-1} , A _{i-1} , K _{i-11} |
| i | -8 | P _i , A _i , K _{i1} |
| $i+1$ | -3 | P _{i+1} , A _{i+1} , K _{i+11} |
| $i+2$ | 2 | P _{i+2} , A _{i+2} , K _{i+21} |
| $i+3$ | 1 | P _{i+3} , A _{i+3} , K _{i+31} |
| $i+4$ | 0 | P _{i+4} , A _{i+4} , K _{i+41} |
| $i+5$ | 0 | P _{i+5} , A _{i+5} , K _{i+51} |
| <hr/> | | |
| $N-1$ | 2 | P _{N-1} , A _{N-1} , K _{N-11} |
| N | 3 | P _N , A _N , K _{N1} |

第 9 図

特開昭63-199399 (18)



第 10 図